Docket No. 245449US-2

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Trygve Frederik MARTON et al. GAU:

SERIAL NO: NEW APPLICATION EXAMINER:

FILED: HEREWITH

FOR: ECHO CANCELLER WITH REDUCED REQUIREMENT FOR PROCESSING POWER

REQUEST FOR PRIORITY

COMMISSIONER FOR PATENTS
ALEXANDRIA, VIRGINIA 22313

SIR:

☐ Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of **35 U.S.C. §120**.

☐ Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of **35 U.S.C. §119(e)**: **Application No.** **Date Filed**

■ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of **35 U.S.C. §119**, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

COUNTRY	**APPLICATION NUMBER**	**MONTH/DAY/YEAR**
Norway	20031103	MARCH 10, 2003

Certified copies of the corresponding Convention Application(s)

- ■ is submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number
  Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and
- ☐ (B) Application Serial No.(s)
  - ☐ are submitted herewith
  - ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Gregory J. Maier
Registration No. 25,599

David A. Bilodeau
Registration No. 42,325

Customer Number
22850
Tel. (703) 413-3000
Fax. (703) 413-2220
(OSMMN 05/03)

KONGERIKET NORGE
The Kingdom of Norway

Bekreftelse på patentsøknad nr
Certification of patent application no

2003 1103

Det bekreftes herved at vedheftede dokument er nøyaktig utskrift/kopi av ovennevnte søknad, som opprinnelig inngitt 2003.03.10

It is hereby certified that the annexed document is a true copy of the above-mentioned application, as originally filed on 2003.03.10

2003.11.04

Line Reum

Line Reum
Saksbehandler

PATENTSTYRET®
Styret for det industrielle rettsvern

BL. 130 - 5.000 - 01.2001 - Hegro Grafisk AS

1d

10. mars 2003

o: 154346

HT/hmm

TP03002NO

SØKER:

Tandberg Telecom AS

Philip Pedersens v. 22

1324 Lysaker

OPPFINNERE:

Trygve Frederik Marton,

Hoff Terrasse 19,

0275 Oslo

Ingvar Flaten Aarnes,

Stolmakergata 9B,

0551 Oslo

TITTEL:

Ekkokansellator med redusert beregningskompleksitet

FULLMEKTIG:

Oslo Patentkontor AS, Postboks 7007M, 0306 Oslo

Field of the invention

The present invention relates to an audio communication system and method with improved acoustic characteristics, and particularly to a video conferencing system including an improved audio echo cancellation system.

Background of the invention

In a conventional conferencing system set-up that uses loudspeakers, two or more communication units are placed at separate sites. A signal transmitted from one site to another site using a conference system experiences several delays, these delays will include a transmission delay and a processing delay. For a video conferencing system, the processing delay for video signals is considerably larger than the processing delay for the audio signals. Because the video and audio signals have to be presented simultaneously, in phase, a lip sync delay is purposefully introduced to the audio signal, in both the transmitting and receiving signal paths in order to compensate for the longer video signal delay.

In a conventional conferencing system, one or more microphones captures a sound wave at a site A, and transforms the sound wave into a first audio signal. The first audio signal is transmitted to a site B, where a television set or an amplifier and loudspeaker, reproduces the original sound wave by converting the first audio signal generated at site A into the sound wave. The produced sound wave at site B, is captured partially by the audio capturing system at site B, converted to a second audio signal, and transmitted back to the system at site A. This problem of having a sound wave captured at one site, transmitted to another site, and then transmitted back to the initial site is referred to as acoustic echo. In its

most severe manifestation, the acoustic echo might cause feedback sound, when the loop gain exceeds unity. The acoustic echo also causes the participants at both site A and site B to hear themselves, making a conversation over the conferencing system difficult, particularly if there are delays in the system set-up, as is common in video conferencing systems, especially due to the above mentioned lip sync delay. The acoustic echo problem is usually solved using an acoustic echo canceller, described below.

Figure 1 shows a conventional conferencing system set-up. For simplicity, Figure 1 shows the conferencing system set-up distributed at two sites, A and B. The two sites are connected through a transmission channel 1300 and each site has a loudspeaker 1100 and 1200, respectively, and a microphone 1111 and 1211, respectively. The arrows in Figure 1 indicate the direction of propagation for an acoustic signal, usually from the microphone to the loudspeaker.

Figure 2 is an overall view of a video conferencing system. This system is distributed at two sites, A and B. As for the conferencing system set-up, a video conferencing module can be distributed at more than two sites and also the system set-up is functional when only one site has a loudspeaker. The video module has at site A a video capturing system 2141 that captures a video image and a video subsystem 2150 that encodes the video image. In parallel, a sound wave is captured by an audio capturing system 2111 and an audio subsystem 2130 encodes the sound wave to the acoustic signal. Due to processing delays in the video encoding system, the control system 2160 introduces additional delays to the audio signal by use of a lip sync delay 2163 so to achieve synchronization between

the video and audio signals. The video and audio signals are mixed together in a multiplexer 2161 and the resulting signal, the audio-video signal is sent over the transmission channel 2300 to site B. Additional lipsync delay 2262 is inserted at site B. Further, the audio signal presented by the audio presenting device 2221 is materialized as a sound wave at site B. Part of the sound wave presented at site B arrives to the audio capturing device 2211 either as a direct sound wave or as a reflected sound wave. Capturing the sound at site B and transmitting this sound back to site A together with the associated delays forms the echo. All delays described sums up to be considerable and therefore the quality requirements for an echo canceller in the video conferencing system are particularly high.

Fig. 3 shows an example of an acoustic echo canceller subsystem, which may be a part of the audio system in the video conferencing system of figure 2. At least one of the participant sites has the acoustic echo canceller subsystem in order to reduce the echo in the communication system. The acoustic echo canceller subsystem 3100 is a full band model of a digital acoustic echo canceller. A full band model processes a complete audio band (e.g., up to 20 kHz; for video conferencing the band is typically up to 7 kHz, in audio conferencing the band is up to 3.4 kHz) of the audio signals directly.

As already mentioned, compensation of acoustic echo is normally achieved by an acoustic echo canceller. The acoustic echo canceller is a stand-alone device or an integrated part in the case of the communication system. The acoustic echo canceller transforms the acoustic signal transmitted from site A to site B, for example, using a

linear/non-linear mathematical model and then subtracts the mathematically modulated acoustic signal from the acoustic signal transmitted from site B to site A. In more detail, referring for example to the acoustic echo canceller subsystem 3100 at site B, the acoustic echo canceller passes the first acoustic signal 3131 from site A through the mathematical modeller of the acoustic system 3121, calculates an estimate 3133 of the echo signal, subtracts the estimated echo signal from the second audio signal 3132 captured at site B, and transmits back the second audio signal 3135, less the estimated echo to site A. The echo canceller subsystem of figure 3 also includes an estimation error, i.e., a difference between the estimated echo and the actual echo, to update or adapt the mathematical model to a background noise and changes of the environment, at a position where the sound is captured by the audio capturing device.

The model of the acoustic system 3121 used in most echo cancellers is a FIR (Finite Impulse Response) filter, approximating the transfer function of the direct sound and most of the reflections in the room. The FIR filter will preferably not, mainly due to processing power, provide echo cancellation in an infinite time after the signal was captured by the loudspeaker. Instead, it will accept that the echo after a given time, the so-called tail length, will not be cancelled, but will appear as residual echo.

To estimate the echo in the complete tail length, the FIR filter will need a length L = Fs * tail length, where Fs is the sampling frequency in Hz, and where the tail length is given in seconds.

The required number of each of the multiplications and additions to calculate one single sample output of the filter equals the filter length, and the output of the filter should be calculated once per sample. I.e. the total number of multiplications and additions are Fs*L = Fs * Fs* tail length = tail length*Fs^2.

A typical value for a tail length is 0.25 sec. The number of multiplications and additions for Fs= 8 kHz system will be 16 Millions, for 16 kHz 64 Millions and for 48 kHz 576 Millions.

Similar calculations could be performed for the filter update algorithm. The simplest algorithm, LMS (Least Mean Square), has a complexity proportional to the filter length, which implies a processing power requirement proportional to Fs^2, while more complex algorithms have processing power proportional to the square of the filter length, which implies a processing power requirement proportional to Fs^3.

One way of reducing the processing power requirements of an echo canceller is to introduce sub-band processing, i.e. the signal is divided into bands with smaller bandwidth, which can be represented using a lower sampling frequency. An example of such system is illustrated in fig. 4.

Analyse filters 4125, 4131 divides the full band signals from far end and near end, respectively, in N sub-bands. The echo cancellation and miscellaneous sub-band processing (typically, but not limited to non-linear processing and noise reduction) is performed in each sub-band, and thereafter a synthesize filter 5127 recreates the modified full band signals. Note that in the following complexity calculations, many minor processing blocks are omitted, as their contribution to the overall processing power requirements are small.

The analyse filters 4125, 4131 includes a filter bank and a decimator, while the synthesize filter 5127 includes a filter bank and an interpolator. The full band signals have sampling frequency $Fs_{fullband}$. The sub-band signals will have a sampling frequency of $Fs_{sub\text{-}band} = K/N*Fs_{fullband}$. K is an over sampling factor, introduced to simplify and reduce the processing power requirements of the filter bank. K is always larger than one, but most often relatively small, typically less than two.

The processing power for the filtering and adaptation (assuming FIR and LMS) for the sub-band case is: $O_{sub\text{-}band} = c_1*taillength*\ Fs_{sub\text{-}band}^2 = c_1*tail\ length*(\ K/N*Fs_{fullband})^2$ (c_1 is a proportionally constant). Thus, for a high N, the processing power requirements of the filtering can be reduced, however, for the total processing power, the overhead of the analyse and synthesize filters must be added.

Effective methods of analyzing and synthesizing the signals are based on a transform, for example a FFT. The methods have complexity $O_{overhead} = c_2*N*\ \log_2 N$, where N is the number of subbands, and c_2 is a proportionally constant. The number of subbands will be proportional with $Fs_{fullband}$, and thus $O_{overhead} = c_3*Fs_{fullband}*\log_2\ Fs_{fullband}$.

I.e. the total complexity is:

$$O = O_{subband} + O_{overhead} = c_1*taillength*(\ K/N*Fs_{fullband})^2 + c_3*Fs_{fullband}*\log_2\ Fs_{fullband}.$$

The echo filtering/adaption is proportional to $Fs_{fullband}^2$. It is possible to reduce the filtering/adaption part by increasing the number of subbands, but at the expense of increased overhead for the calculations of the subband signals. Still, by using a large number of subband, i.e.

using a large fast transform, it is possible to obtain a complexity which increases with $Fs_{fullband} \log_2 Fs_{fullband}$.

Though theoretically possible, this may be difficult to achieve in practical implementations, due to cache ineffiency in signal processing when applying large transforms.

Thus, efforts have been made for providing a system allowing reduction in the number of sub-bands without increasing the sub-bandwidths.

Summary of the invention

It is an object of the present invention to provide a system allowing reduction in the number of sub-bands without increasing the sub-bandwidths.

The features defined in the independent claim enclosed characterise this system.

In particular, the present invention discloses an audio echo canceller comprising a module at least configured to implement a model of acoustic echo for thereby providing an echo estimate and subtracting the estimate from a first signal derived from a first decimator, in addition to an interpolator configured to interpolate a second signal derived from the module, which canceller includes a subtracting device adapted to provide said second signal by subtracting the output of said module with the first signal being adjusted by a first filter, and an adding device adapted to add the input signal of said decimator being adjusted by a second filter, to the output signal of said interpolator.

Brief d scription of the drawings

In order to make the invention more readily understandable, the discussion that follows will refer to the accompanying drawings,

Figure 1 is an overview block diagram of a conventional conferencing system set-up,

Figure 2 is a more detailed block diagram of a conventional conferencing system set-up,

Figure 3 is a closer view of an acoustic echo canceller subsystem,

Figure 4 is a block diagram of the corresponding echo canceller subsystem implemented with sub-band processing,

Figure 5 is a block diagram of an echo canceller subsystem implemented with sub-band processing according to the present invention.

Best mode of carrying out the invention

In the following, the present invention will be discussed by describing a preferred embodiment, and by referring to the accompanying drawings. However, even if the specific embodiment is described in connection with video conferencing, people skilled in the art will realize other applications and modifications within the scope of the invention as defined in the enclosed independent claim.

The present invention takes advantage in the fact that not all frequencies are equally important in a high frequency echo cancelling system.

Frequencies above approximate 7 kHz does not contribute much to the speech intelligibility. However, these

frequencies impacts on the naturalness and experienced vicinity are considerable.

Experience has shown that both speech intelligibility and listening impression are maintained even when the returning signal at certain occurrences is being low pass filtered and down sampled. In other words, an audio echo cancellation system may advantageously be designed having:

a) full duplex communication (including echo cancelling) in frequencies which contributes to speech intelligibility, to ensure that no information is lost during double talk, and

b) full bandwidth and increased naturalness during periods with single talk.

By embodying the above limitations properly, the exhaustive sampling frequency influence on the processing power requirements may be reduced, while still obtaining the benefits of full audible bandwidth sound.

The present invention provides a system wherein echo cancellation and noise reduction is treated as in prior art in communication critical frequency bands, while above this limit, voice switching is preferably used to provide high fidelity speech, and at the same time avoiding echo and feedback.

Figure 5 shows a preferred embodiment of the present invention. It is based upon the sub-band echo cancellation system of figure 4. The overall system of the preferred embodiment operates at samplerate Fs_{high}, and the echocanceller working on samplerate Fs_{low}, processing sound with frequencies below Fs_{low} /2. Note that the processing box 5000 is repeated for all sub-bands.

Before being processed by the echocanceller the signal from site B, including echo, near end sound and/or noise is decimated, i.e. lowpassfiltered and downsampled by a factor n. The signal is also tapped and forwarded for further processing, and constitute the part of the output signal with high frequencies (above $Fs_{low}/2$). The lowpassfiltered and downsampled signal is divided into N sub-bands by the analyse filter. Since the signal that is to be divided in the preferred embodiment of the present invention is lowpass filtered, the required magnitude of N will be reduced correspondingly.

The sub-band signal 5132 is then added to an inverted sub-band echo estimate 5133 being generated by a (sub-)model 5121 of the acoustic system. As in prior art, the model includes preferably a FIR filter and an associated filter update algorithm, e.g. a LMS algorithm, having the corresponding sub-band signal of the audio signal from site A and a feedback loop from the result of the above-mentioned addition 5134 as inputs. The resulting signal 5134 is preferrably further processed by miscellaneous processing, e.g residual echo masking (due to the finite nature of the FIR filter, and any other model infirmities), noise reduction and comfort noise addition. The resulting signal after the miscellaneous processing will comprise the noise reduced and echo free sub-band signal from site B in addition to comfort noise.

The above mentioned tapped signal should in some way be high pass filtered as it intends to contribute the high frequency part of the output signal. According to the present invention, this is achieved simply by subtracting the low-pass filtered signal from the original signal. The low pass filtered signal could be provided by tapping it right after the lowpassfilter in the decimator (delaying the mic signal by the proper amount of samples T), but this is not preferrable as it would make the decimator processing less efficient by prohibiting integration of

lowpassfiltering and downsampling. The preferred way is to subtract the clean sub-band signal tapped right after the analyse filter from the processed sub-band signal right before the synthese filter. This will make a path in the sub-band processing part merely providing the low frequencies of the site B signal, which could be used for high pass filtering the by-passing signal by means of the already mentioned subtraction.

The lowpass filters 5142, 5139, the downsampler 5141 and upsampler 5140 governs the highpass filter's profile, together with H_s 5136 and H_f 5138, which are further explained below. Having appropriate delay of the bypassing signal is of course crucial for this kind of filtering, this delay must be added both before and after the filtering by H_f, as H_f's magnitude must correspond to H_s. These delays should represent the delays in the lowpass-, analysis- and synthesis-filters, as well as any additional delays.

The present invention provides echocancelling and noisereduction at low frequencies, and unaltered microphone sound at high frequencies. This is desirable in case of near end talk, i.e. speech at site B, with a minimum of noise. However, without any level adjustments it might produce feedback, and high frequency echoes will pass right through. Feedback may even damage hearing. Hence, it is necessary to identify situations where full audible bandwidth sound is required, and situations where high frequencies should be attenuated, respectively. According to the preferred embodiment of the present invention, the control algorithm 5137 identifies these situations, typically based on (but not limited to) fullband loudspeaker and microphone signal, subband signals of the same, subband echo estimate and echo cancelled subband signal.

The control algorithm 5137 should at least provide information determining the following situations: a) near end talk, or b) either far end talk, double talk or background noise only. In case b), the high frequency part of the sound should then be attenuated by adjusting the gain values of H_s and H_f closely to zero. Far end talk will produce echo, and only noise will contain high frequency components, and might trigger feedback. Double talk situations can be handled without high frequency sound since the ear is less sensitive to high fidelity in sound reproduction while the soundscape is chaotic as for instance when people talk at the same time.

The decisionmaking control algorithm constantly produces values, e.g. 1 for situation a) and 0 for b), which is interpreted and used as basis for the level of high frequency sound, or more precisely, the filters H_s and H_f adjusting the high-pass filter's profile. This is the most important function of the filters H_s and H_f, namely the adjustment of the magnitude of the high frequencies. In the following an example of a highband gain change interpretation of the decision (i.e. transition from a) to b) or vice versa), is disclosed.

Assume the sound is partitioned in packages of duration 10ms, at a samplerate of Fs_{high} = 48000kS/s, and n=3 so that Fs_{low} = 16000kS/s. When the echocanceller operates on sub-bands, the decision is typically taken once a packet, so the adjustment of the filters can be done at most once each sound-packet. As the sub-band samples are representing 10ms each in a narrow frequency band, while the signal samples from site B on the other hand each represents a small amount of time in a relatively broad frequency band, it would be obvious for a man skilled in the art that the result of gain adjustments in the lower frequencies (provided by the magnitude of H_s) should be accordingly treated at high frequencies (i.e. the corresponding magnitude of H_f).

If the decimator, analyze filter, synthesize filter and interpolator together form a linear phase system, H_s can be reduced to time invariant gain G_s while H_f can be replaced by appropriate delays and a time invariant gain G_f. Any change in G_s must be reflected by a timely distributed changing G_f. Only small errors are introduced by calculating G_f as a linear interpolation of consecutive G_s values.

In stable situations, i.e. silence (only noise) or far end talk, the gain of the filters H_s and H_f should stabilize at zero (no high frequency sound/noise). Near end talk will not be a very stable situation, as speech includes both sound (phrases) and silence (between phrases). Still, during phrases it is preferrable to let the gain of H_s and H_f (the maximum high frequency gain) be as constant as possible and equal 1, for thereby producing full audible bandwidth sound.

Note that there might be cases of large amounts of background noise where a lower level of maximum high frequency gain (between 0 and 1) is preferable. The reduction of maximum high frequency gain will of course deteriorate the functionality of the device and should be solved by reducing high frequency background noise when possible.

Even if the present invention is described in connection with videoconferencing, a person skilled in the art would realize that it would also be useful in other equivalent applications like telephone conferences and calls, mobile telephone conferences and calls, web conferences etc.

The main advantage of the present invention is that it requires lower processing power than prior art because of fewer sub-bands. Its complexity is $O = c_4 * Fs_{low} * \log_2 Fs_{low} + c_5 * Fs_{high}$. I.e. when the bandwidth increases above the

communication critical frequency bands, the complexity only scales linearly with the bandwidth.

Further, the system of the present invention may be added as a framework around existing echo cancellers, with none or only minor adjustments in the existing canceller. Thus, the present invention provides an efficient (in terms of development resources) way of increasing the bandwidth of existing echo canceller systems. It can be used with both sub-band and full-band echo cancellers.

In addition, improved audio quality for the near end signal during single talk may be provided. The near end signal transmitted to the far end site has not been passed through the analyse/synthesize (in the sub-band case) filter process, as the magnitude of H_f is 1. Therefore, any distortion or other quality degradations in this process are not added to the near end signal.

P a t e n t   c l a i m s

1. An audio echo canceller comprising a module at least configured to implement a model of acoustic echo for thereby providing an echo estimate and subtracting the estimate from a first signal derived from a first decimator, in addition to an interpolator configured to interpolate a second signal derived from the module, c h a r a c t e r i z e d   i n

a subtracting device adapted to provide said second signal by subtracting the output of said module with the first signal adjusted by a first filter,

an adding device adapted to add the input signal of said decimator adjusted by a second filter, to the output signal of said interpolator.

2. Audio echo canceller according to claim 1, c h a r a c t e r i z e d   i n   a first analyse filter adapted to divide the output of said decimator into a number of first signals of respective sub frequency bands, and a synthesize filter adapted to combine said number of second signals to an input of said interpolator.

3. Audio echo canceller according to claim 2, c h a r a c t e r i z e d   i n   a second analyse filter adapted to divide the output of a second decimator into said number of third signals of respective sub frequency bands, one of which being an input to said model of the acoustic wave.

4. Audio echo canceller according to claim 3, c h a r a c t e r i z e d   i n   a control module adapted to adjust the gain of said first and second filter dependent on the first and the third signal.

5. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that the first and the second decimator both include a low pass filter and a down sampler, and the interpolator includes an up sampler and a low pass filter.

6. Audio echo canceller according to claim 5,
c h a r a c t e r i z e d i n that one or more of the low pass filters are implemented as FIR filters.

7. Audio echo canceller according to claim 5 or 6,
c h a r a c t e r i z e d i n that the analyse filter, the synthesize filter and one or more of the low pass filters are linear phase.

8. Audio echo canceller according to claim 7,
c h a r a c t e r i z e d i n that the first and the second filters are time variant amplifiers.

9. Audio echo canceller according to one of the claims 1 - 6,
c h a r a c t e r i z e d i n that the first and the second filters are amplifiers.

10. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that said model includes a FIR filter and an associated filter update algorithm.

11. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that said module further includes a miscellaneous processing unit, at least including residual echo masker, a noise reduction algorithm and a comfort noise generator.

12. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that the canceller is a part of a video conferencing system wherein the input of the second decimator is a second audio signal captured by a microphone at a far end site including far end sound, and the input of the first decimator is a first audio signal captured by a microphone at a near end site including near end sound, noise and/or the acoustic echo.

13. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that the canceller is a part of a telephone communication and/or conferencing system wherein the input of the second decimator is a second audio signal captured by a microphone at a far end site including far end sound, and the input of the first decimator is a first audio signal captured by a microphone at a near end site including near end sound, noise and/or the acoustic echo.

14. Audio echo canceller according to one of the preceding claims,
c h a r a c t e r i z e d i n that the canceller is a part of a mobile communication and/or conferencing system wherein the input of the second decimator is a second audio signal captured by a microphone at a far end site including far end sound, and the input of the first decimator is a first audio signal captured by a microphone at a near end site including near end sound, noise and/or the acoustic echo.

15. Audio echo canceller according to claim 12, 13 or 14,
c h a r a c t e r i z e d i n that the control module is adapted to detect the presence and/or content of said first and second audio signal and adjust said gain accordingly.

16. Audio echo canceller according to claim 15, c h a r a c t e r i z e d i n that the control module is adapted to adjust the gain to a first non-zero value, preferably one, if near end sound together with noise, or near end sound only is detected, and to a second zero value in all other cases.

17. Audio echo canceller according to one of the preceding claims, c h a r a c t e r i z e d i n one or more delay units before and/or after the second filter, which accumulated correspond to a delay appearing from the first decimator to the interpolator.

18. A method in an audio echo cancellation system comprising a module at least configured to implement a model of acoustic echo for thereby providing an echo estimate and subtracting the estimate from a first signal derived from a first decimator, in addition to an interpolator configured to interpolate a second signal derived from the module, c h a r a c t e r i z e d i n

tapping the first signal before said module,

adjusting the tapped first signal by a first filter,

providing said second signal by subtracting the output of said module with the adjusted first signal,

tapping the input signal of the decimator,

adjusting the tapped input signal of the decimator by a second filter,

adding the adjusted input signal of said decimator to the output signal of said interpolator.

19. Method according to claim 18,
c h a r a c t e r i z e d i n the following further steps:

dividing the output of said decimator into a number of first signals of respective sub frequency bands, and

providing the input of said interpolator by combining said number of second signals.

20. Method according to claim 19,
c h a r a c t e r i z e d i n the following further step:

dividing the output of a second decimator into said number of third signals of respective sub frequency bands, one of which being an input to said model of the acoustic wave.

21. Method according to claim 20,
c h a r a c t e r i z e d i n the following further step:

adjusting the gain of said first and second filter dependent on the first and the third signal.

22. Method according to one of the claims 18 - 21,
c h a r a c t e r i z e d i n the following further steps:

in the first and the second decimator, low pass filtering and down sampling the input signal of the decimators, respectively,

in the interpolator, low pass filtering and up sampling the input of the interpolator.

23. Use of method according to claim 18 - 22 in a video, telephone or mobile conferencing and/or communication system.

Abstract

The present invention discloses an improved audio echo canceller processing echo, noise and near end talk in a narrower, but still intelligible, frequency band for reducing required processing power and complexity. In a preferred embodiment of the present invention, an input audio signal of captured sound in an audio communication system is decimated and then divided into a number of sub bands by an analyse filter. Each sub band is processed as in conventional audio echo cancelling by subtracting the signal with an echo estimate from a model of the acoustic signal in the respective sub band, except from that the signal is also bypassed, adjusted by an filter and subtracted from the processed signal. The resulting signals are then recombined by a synthesise filter and interpolated to the original sampling rate and bandwidth. Finally, the output from the synthesize filter is added to the input audio signal, which has been delayed and adjusted by an filter. The filters are controlled by an control algorithm detecting the presence of near end sound, far end sound and noise, so that the filters, and consequently the high pass filter of the echo canceller, only pass high frequency (above low pass frequencies) in cases where only near end sound is detected.

1p

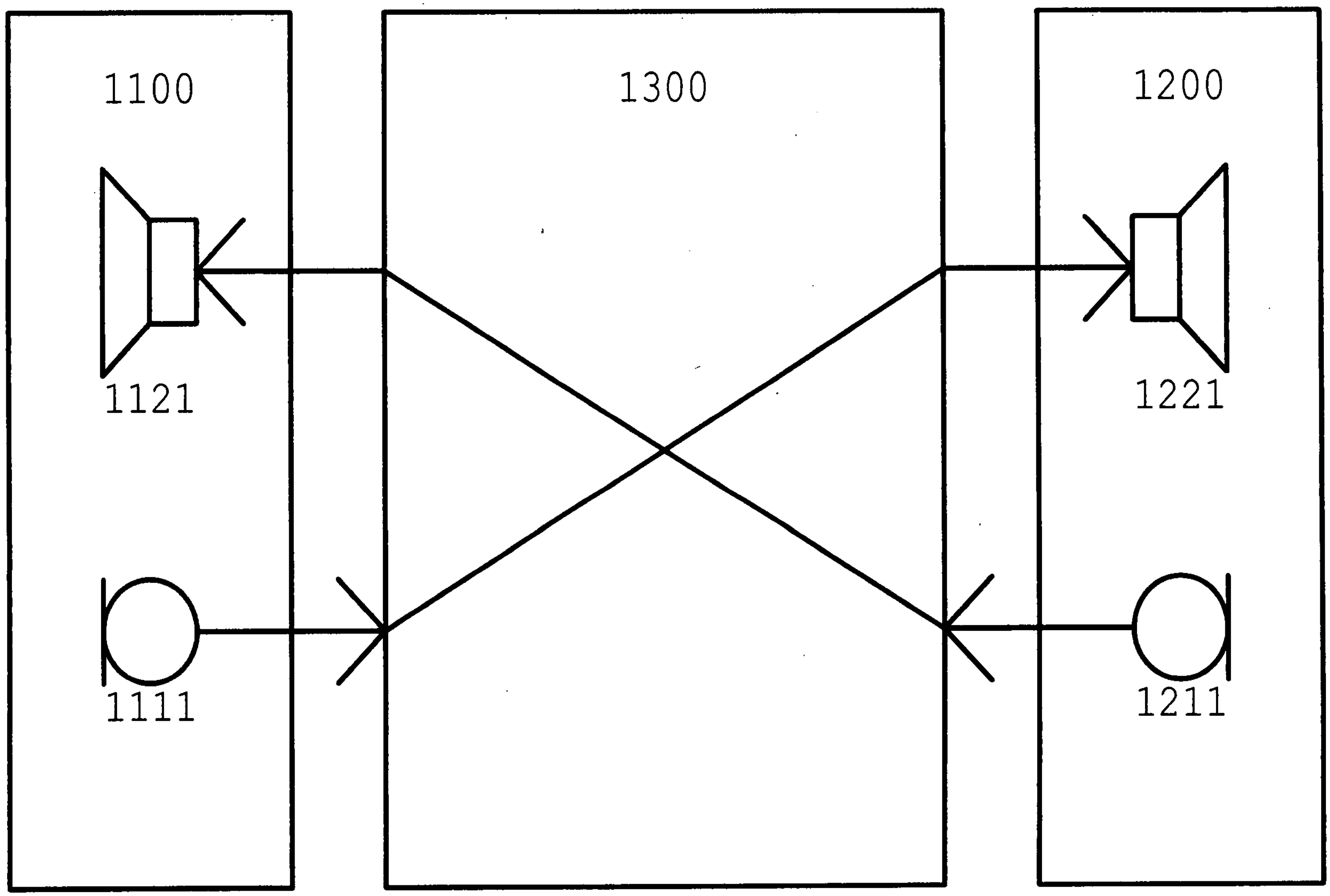

FIGURE 1 (BACKGROUND ART)

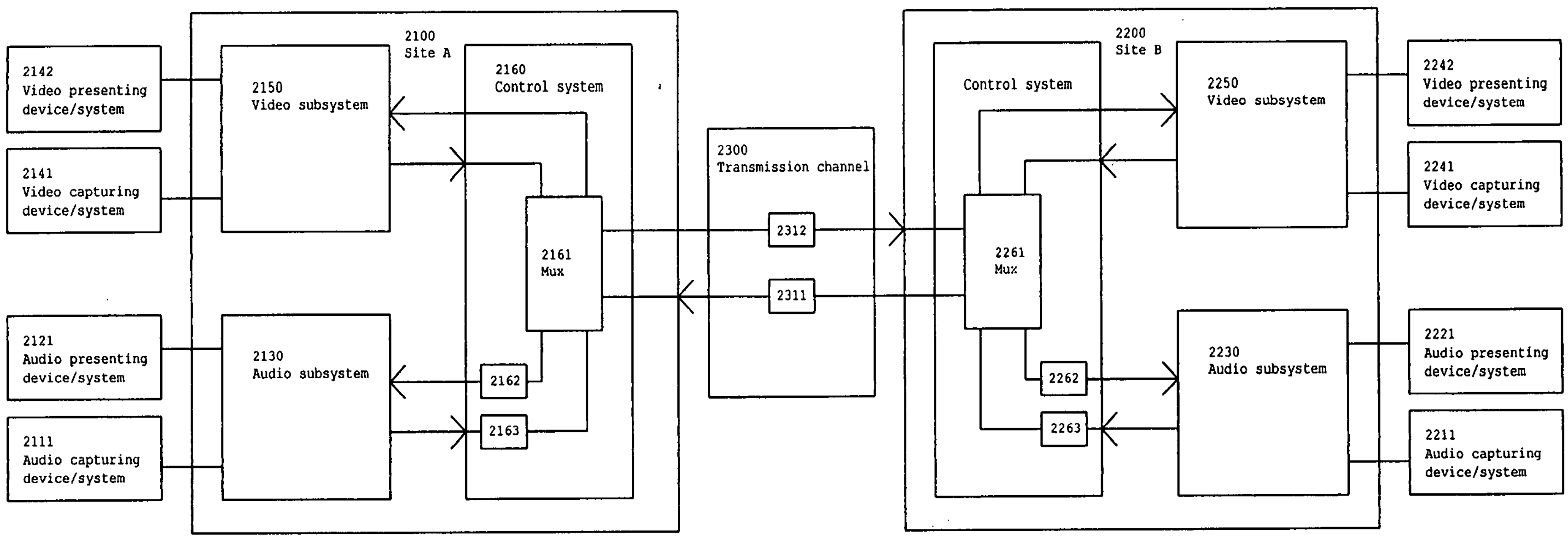

FIGURE 2 (BACKGROUND ART)

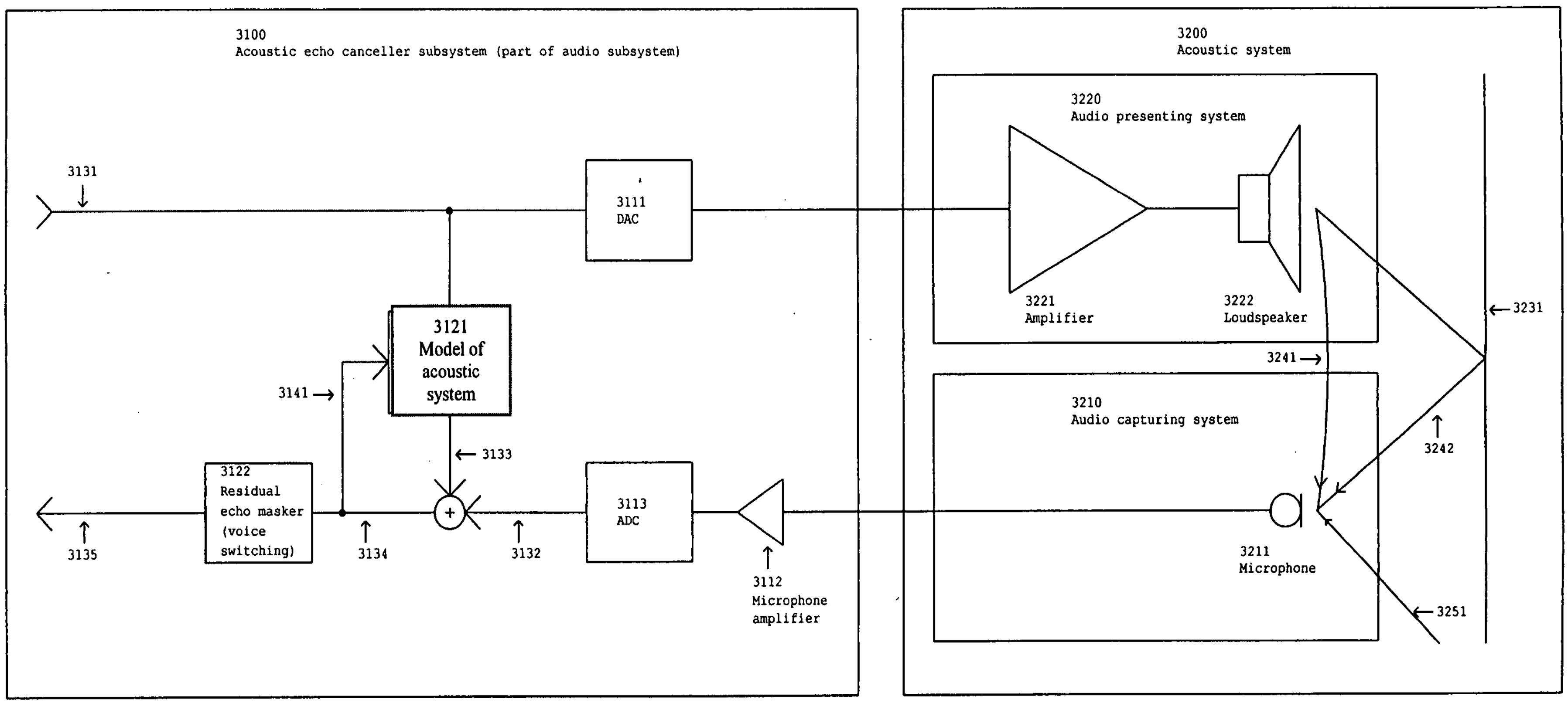

FIGURE 3 (BACKGROUND ART)

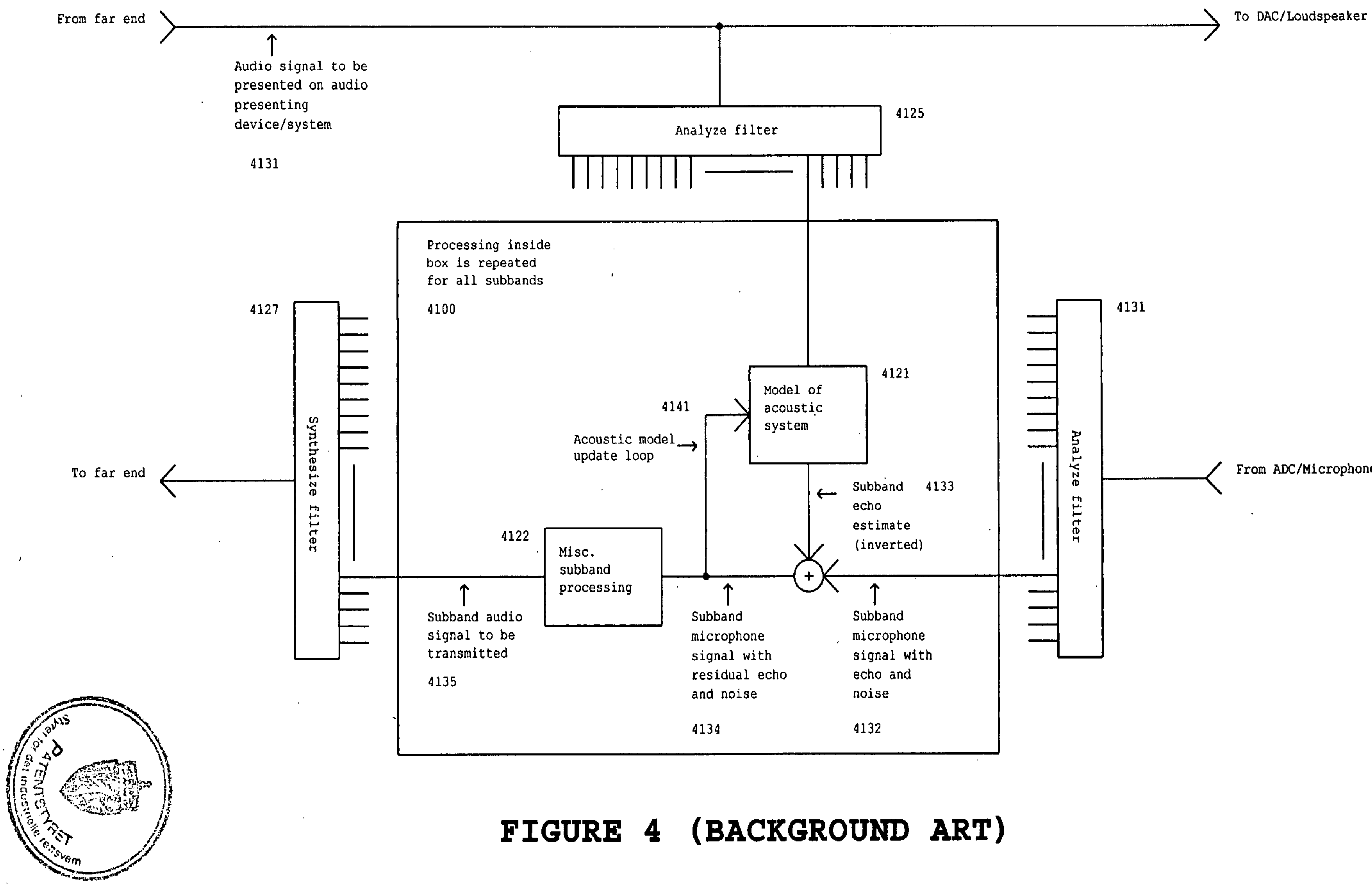

FIGURE 4 (BACKGROUND ART)

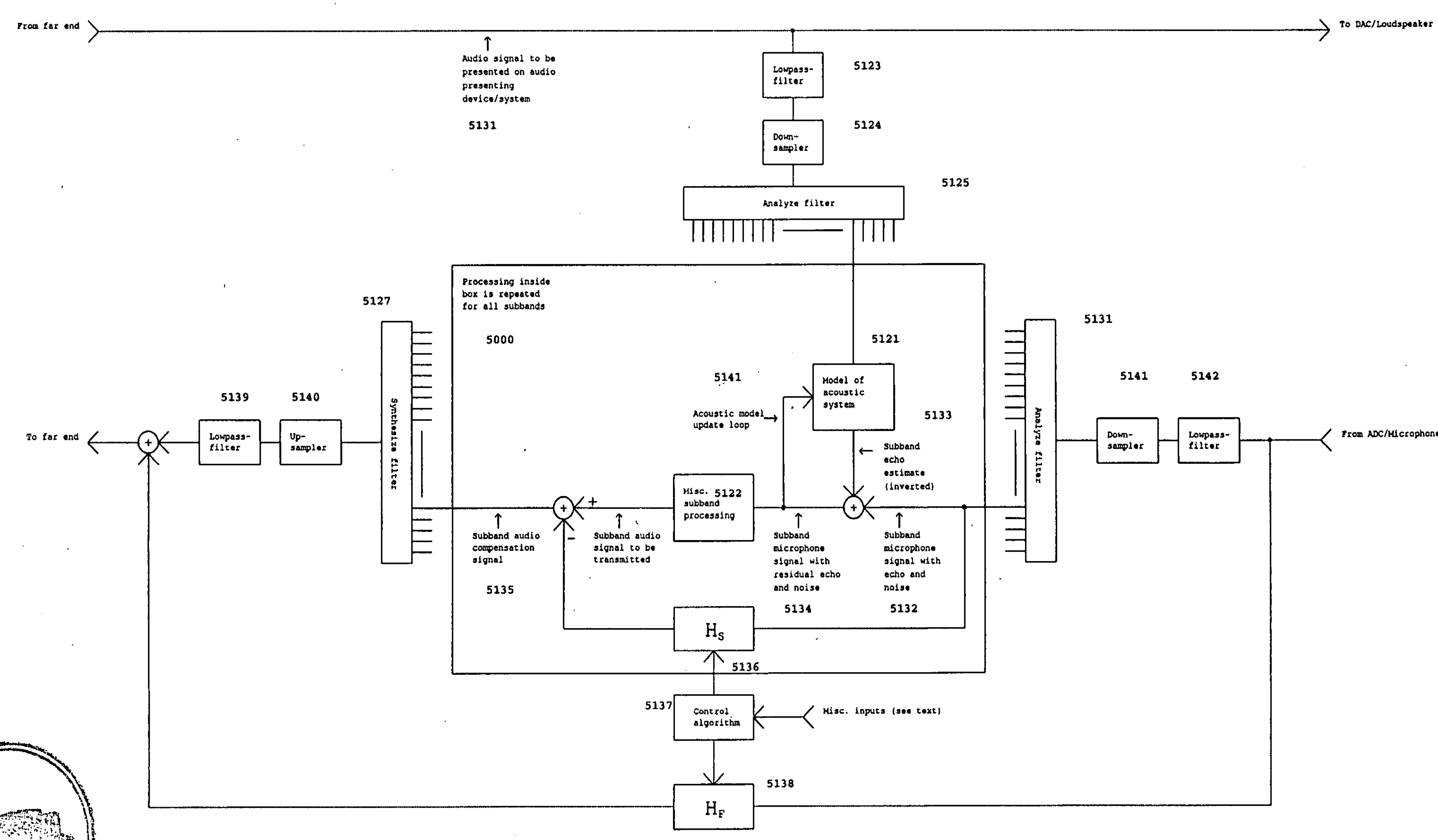
From far end
To DAC/Loudspeaker
Audio signal to be presented on audio presenting device/system
5131
Lowpass-filter
5123
Down-sampler
5124
Analyze filter
5125
Processing inside box is repeated for all subbands
5000
5127
Synthesize filter
5139
5140
Lowpass-filter
Up-sampler
To far end
5121
Model of acoustic system
5141
Acoustic model update loop
5133
Subband echo estimate (inverted)
Misc. 5122 subband processing
Subband audio compensation signal
Subband audio signal to be transmitted
Subband microphone signal with residual echo and noise
Subband microphone signal with echo and noise
5135
5134
5132
H_S
5136
5137
Control algorithm
Misc. inputs (see text)
5138
H_F
5131
Analyze filter
5141
5142
Down-sampler
Lowpass-filter
From ADC/Microphone

FIGURE 5